

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:29 ; Search time 12.89 seconds
(without alignments)
115,590 Million cell updates/sec

Title: US-09-829-481-4

Perfect score: 339

Sequence: 1 MKSTAIIFVLVAFCLDGL.....IRRGFGCGTFTTCVCYR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68.5	20.2	72	4	US-09-003-198A-17
2	66	19.5	40	1	US-08-385-375-16
3	66	19.5	40	1	US-08-385-375-39
4	64.5	19.0	80	4	US-09-103-489-20
5	64	18.9	74	4	US-09-442-631-4
6	63.5	18.7	768	1	US-08-454-455-4
7	63.5	18.7	769	1	US-08-454-455-6
8	61.5	18.1	80	1	US-08-377-687-49
9	61.5	18.1	80	2	US-08-777-192-49
10	61.5	18.1	80	4	US-08-971-982-49
11	60.5	17.8	81	4	US-09-053-021-4
12	60.5	17.8	81	4	US-09-053-021-9
13	60.5	17.8	652	2	US-08-751-305-2
14	60	17.7	79	1	US-08-627-706-15
15	60	17.7	79	4	US-09-103-489-15
16	59.5	17.6	42	2	US-08-751-305-3
17	59.5	17.6	80	1	US-08-377-687-59
18	59.5	17.6	80	2	US-08-777-192-59
19	59.5	17.6	80	4	US-08-971-982-59
20	59	17.4	204	1	US-08-652-859-2
21	59	17.4	204	2	US-08-919-706-2
22	59	17.4	204	2	US-09-153-751-2
23	58.5	17.3	75	1	US-08-289-458-2
24	58.5	17.3	75	2	US-08-761-549-2
25	58.5	17.3	75	4	US-09-127-646-2
26	57.5	17.0	221	4	US-08-925-708-1
27	57	16.8	86	1	US-08-149-839B-14

28	57	16.8	86	1	US-08-451-568-14	Sequence 14, Appl
29	57	16.8	86	1	US-08-451-566-14	Sequence 14, Appl
30	57	16.8	86	2	US-08-777-113-14	Sequence 14, Appl
31	57	16.8	1394	6	5177197-30	Patent No. 5177197
32	57	16.8	1400	4	US-08-630-915A-37	Sequence 37, Appl
33	56.5	16.7	80	3	US-08-952-383A-14	Sequence 14, Appl
34	56.5	16.7	1248	4	US-08-882-046-6	Sequence 6, Appl
35	56	16.5	63	1	US-08-117-080-12	Sequence 12, Appl
36	56	16.5	63	1	US-08-471-329-12	Sequence 12, Appl
37	56	16.5	63	2	US-08-915-142-12	Sequence 12, Appl
38	56	16.5	67	4	US-09-053-021-2	Sequence 2, Appl
39	56	16.5	89	1	US-08-181-556-2	Sequence 2, Appl
40	56	16.5	103	1	US-08-486-013-20	Sequence 20, Appl
41	56	16.5	103	2	US-08-482-279-20	Sequence 20, Appl
42	56	16.5	103	2	US-08-342-268-20	Sequence 20, Appl
43	56	16.5	103	3	US-09-015-968-20	Sequence 20, Appl
44	56	16.5	103	4	US-09-397-386-20	Sequence 20, Appl
45	55.5	16.4	47	1	US-08-656-318A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-003-198A-17
; Sequence 17, Application US/09003198A
; Patent No. 6316407
; GENERAL INFORMATION:
; APPLICANT: Liang, Jihong
; APPLICANT: Shah, Dilip Maganlal
; APPLICANT: Wu, Yonnie S.
; APPLICANT: Rosenberger, Cindy A.
; APPLICANT: Hakimi, Salim
; TITLE OF INVENTION: Antifungal Polypeptide and Methods for
; TITLE OF INVENTION: Controlling Plant Pathogenic Fungi
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/003.198A
; FILING DATE: 07-JAN-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Patterson, Melinda L.
; REGISTRATION NUMBER: 33,062
; REFERENCE/DOCKET NUMBER: MOBT:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 787-1400
; TELEFAX: (713) 787-1440
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-003-198A-17

Query Match 20.2%; Score 68.5; DB 4; Length 72;
Best Local Similarity 28.4%; Pred No. 0.26;
Matches 19; Conservative 11; Mismatches 24; Indels 13; Gaps 3;
Oy 2 KSIAIIFVLVAFCLDGLVEAGFGCPFNACK-----CHRHCKSIIRRG---GFCR 50

Db 4 KSLACLFLLLVFAQELVSEANTCENLAGYKGVCFGGCDRHCR--QEGAI--GRGR 61
QY 51 GTFRTC 57
Db 62 DDFRCWC 68

RESULT 2
US-08-385-375-16
; Sequence 16, Application US/08385375
; Patent No. 5631144
; GENERAL INFORMATION:
; APPLICANT: LEMOINE, Yves
; APPLICANT: NGUYEN, Martine
; APPLICANT: ACHSTETTER, Tilman
; APPLICANT: REICHART, Jean-Marc
; TITLE OF INVENTION: APPLICATION OF NOVEL DNA FRAGMENTS AS A
; TITLE OF INVENTION: SEQUENCE CODING FOR A SIGNAL PEPTIDE FOR THE SECRETION OF
; TITLE OF INVENTION: MATURE PROTEINS BY RECOMINANT YEASTS, EXPRESSION
; TITLE OF INVENTION: CASSETTES, TRANSFORMED YEASTS AND CORRESPONDING PROCESS FOR
; TITLE OF INVENTION: THE PREPARATION OF PROTEINS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385.375
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,356
; FILING DATE: 04-JAN-1994
; APPLICATION NUMBER: FR 89/05687
; FILING DATE: 28-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR90/00306
; FILING DATE: 27-APR-1990
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-385-375-16

Query Match 19.5%; Score 66; DB 1; Length 40;
Best Local Similarity 39.4%; Pred. No. 0.28;
Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;
QY 27 GCPFNAGKCHRHCKSIIRRGCGFRGTFRTCVC 59
Db 8 GTGINHSACAAHCLLRGNRGYCNG--KGVCVC 38

RESULT 3
US-08-385-375-39
; Sequence 39, Application US/08385375
; Patent No. 5631144
; GENERAL INFORMATION:
; APPLICANT: LEMOINE, Yves
; APPLICANT: NGUYEN, Martine
; APPLICANT: ACHSTETTER, Tilman
; APPLICANT: REICHART, Jean-Marc
; TITLE OF INVENTION: APPLICATION OF NOVEL DNA FRAGMENTS AS A
; TITLE OF INVENTION: SEQUENCE CODING FOR A SIGNAL PEPTIDE FOR THE SECRETION OF
; TITLE OF INVENTION: MATURE PROTEINS BY RECOMINANT YEASTS, EXPRESSION
; TITLE OF INVENTION: CASSETTES, TRANSFORMED YEASTS AND CORRESPONDING PROCESS FOR
; TITLE OF INVENTION: THE PREPARATION OF PROTEINS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385.375
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,356
; FILING DATE: 04-JAN-1994
; APPLICATION NUMBER: FR 89/05687
; FILING DATE: 28-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR90/00306
; FILING DATE: 27-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-385-375-39

Query Match 19.5%; Score 66; DB 1; Length 40;
Best Local Similarity 39.4%; Pred. No. 0.28;
Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;
QY 27 GCPFNAGKCHRHCKSIIRRGCGFRGTFRTCVC 59
Db 8 GTGINHSACAAHCLLRGNRGYCNG--KGVCVC 38

RESULT 4
US-09-103-489-20
; Sequence 20, Application US/09103489
; Patent No. 6215048
; GENERAL INFORMATION:
; APPLICANT: Liang, Jihongq
; APPLICANT: Shah, Dilip M.
; APPLICANT: Wu, Yonnie S.

```

RESULT      5
US-09-442-631-4
; Sequence 4, Application US/09442631
; Patent No. 6300489
; GENERAL INFORMATION:
; APPLICANT: OH, BOUNG-JUN
; APPLICANT: KO, MOON KYUNG
; APPLICANT: SHIN, BYONGCHUL
; APPLICANT: CHUNG, CHANG HO
; TITLE OF INVENTION: SMALL AND CYSTEINE RICH ANTIFUNGAL DEFENSIN AND
; TITLE OF INVENTION: THIONIN-LIKE PROTEIN GENES HIGHLY EXPRESSED IN THE
; TITLE OF INVENTION: INCOMPATIBLE INTERACTION
; FILE REFERENCE: 1942/44
; CURRENT APPLICATION NUMBER: US/09/442,631
; CURRENT FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Capsicum annuum
US-09-442-631-4

```

Query Match	18.78;	Score	63.5;	DB	1;	Length	768;
Best Local Similarity	24.78;	Pred. No.	10;				
Matches	19;	Conservative	8;	Mismatches	11;	Indels	39;
Gaps							
Qy	14	PCILEDGIVEAGFCGPFN-----AGKCH-----				RRCKSTR	44
		:	:	:	:	:	:
Db	545	YCEKDD-----FSCPYYHGLCAGHGECEAGRCQCFSGWEGDRQCQPSAAQHC--VNS					596
Qy	45	RGGFCRGFTRTTCVYR					61


```
;
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/777.192
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002.480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-777-192-49

Query Match 18.1%; Score 61.5; DB 2; Length 80;
Best Local Similarity 30.0%; Pred. No. 1.9;
Matches 21; Conservative 8; Mismatches 28; Indels 13; Gaps 4;

QY 4 IAIIFVLVAFCLILE-DGIVEAGFGCFPNAG-----KCHRHCKSIIR-RGGFCRG 51
||:| || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8 IALLFAALVLEAFAPTMTVEAQKLCRPSGTWSGVCGNNACKNQCNLEKARHGSCNY 67

-QY 52 TFRT-TCVCY 60
| | | | |
Db 68 VFPAHKCICY 77

RESULT 10
US-08-971-982-49
; Sequence 49, Application US/08971982
; Patent No. 6187904
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; CAMMUE, BRUNO P.A.
; OSBORN, RUPERT W.
; REES, SARAH B.
; TERRAS, FRANKY R.G.
; VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,982
; FILING DATE: 17-NO. 6187904-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US 08/002.480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-971-982-49

Query Match 18.1%; Score 61.5; DB 4; Length 80;
Best Local Similarity 30.0%; Pred. No. 1.9;
Matches 21; Conservative 8; Mismatches 28; Indels 13; Gaps 4;

QY 4 IAIIFVLVAFCLILE-DGIVEAGFGCFPNAG-----KCHRHCKSIIR-RGGFCRG 51
||:| || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8 IALLFAALVLEAFAPTMTVEAQKLCRPSGTWSGVCGNNACKNQCNLEKARHGSCNY 67

QY 52 TFRT-TCVCY 60
| | | | |
Db 68 VFPAHKCICY 77

RESULT 11
US-09-053-021-4
; Sequence 4, Application US/09053021
; Patent No. 6270785
; GENERAL INFORMATION:
; APPLICANT: SELISKO, Barbara
; APPLICANT: GARCIA-RODRIGUEZ, Consuelo
; APPLICANT: ZAMUDIO-ZUNIGA, Fernando
; APPLICANT: BECERRIL-LUJAN, Baltazar
; APPLICANT: POSSANI-POSTAY, Lourival D.
; TITLE OF INVENTION: Primary Sequence and cDNA of
; Patent No. 6270785
; TITLE OF INVENTION: Insecticidally Effective Toxins from Scorpions of the
; Genus Centruroides
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; OSBORN, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053.021
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017.007
; FILING DATE: 30-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/848.261
; FILING DATE: 29-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett, Arthur S.
; REGISTRATION NUMBER: 20,338
```

; REFERENCE/DOCKET NUMBER: 06899.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-053-021-4

Query Match 17.8%; Score 60.5; DB 4; Length 81;
Best Local Similarity 23.4%; Pred. No. 2.4;
Matches 15; Conservative 10; Mismatches 28; Indels 11; Gaps 3;
QY 4 IAIIFIVLVAFCILEDGIVEAGFGCPFNA-----GKCHRHCKSIIRRG--GFCRGTFRT 55
Db 1 ITACLVLTGTVCAKEGYLVNKSTGCKYNCLILGENKNCDMECKAKNQGGSYGYC---YKL 57
QY 56 TCVC 59
Db 58 ACWC 61

RESULT 12
US-09-053-021-9
; Sequence 9, Application US/09053021
; Patent No. 6270785
; GENERAL INFORMATION:
; APPLICANT: SELISKO, Barbara
; APPLICANT: GARCIA-RODRIGUEZ, Consuelo
; APPLICANT: ZAMUDIO-ZUNIGA, Fernando
; APPLICANT: BECERRIL-LUJAN, Baltazar
; APPLICANT: POSSANI-POSTAY, Lourival D.
; TITLE OF INVENTION: Primary Sequence and cDNA of
; Patent No. 6270785
; TITLE OF INVENTION: Insecticidally Effective Toxins from Scorpions of the
; TITLE OF INVENTION: Genus Centruroides
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,021
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,007
; FILING DATE: 30-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/848,261
; FILING DATE: 29-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett, Arthur S.
; REGISTRATION NUMBER: 20,338
; REFERENCE/DOCKET NUMBER: 06899.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 81 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-053-021-9

Query Match 17.8%; Score 60.5; DB 4; Length 81;
Best Local Similarity 23.4%; Pred. No. 2.4;
Matches 15; Conservative 10; Mismatches 28; Indels 11; Gaps 3;
QY 4 IAIIFIVLVAFCILEDGIVEAGFGCPFNA-----GKCHRHCKSIIRRG--GFCRGTFRT 55
Db 1 ITACLVLTGTVCAKEGYLVNKSTGCKYNCLILGENKNCDMECKAKNQGGSYGYC---YKL 57
QY 56 TCVC 59
Db 58 ACWC 61

RESULT 13
US-08-751-305-2
; Sequence 2, Application US/08751305
; Patent No. 5965439
; GENERAL INFORMATION:
; APPLICANT: Tenner et al., Andrea J.
; TITLE OF INVENTION: HOST DEFENSE ENHANCEMENT
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,305
; FILING DATE: 18-NOV-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07306/012001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-751-305-2

Query Match 17.8%; Score 60.5; DB 2; Length 652;
Best Local Similarity 36.2%; Pred. No. 19;
Matches 17; Conservative 2; Mismatches 13; Indels 15; Gaps 3;
QY 22 VFAGFGCPFNAGKCHRHCKSIIRRG-----FCRGTFR-----TTC 57
Db 258 VSPKYGCNENNGGCHQDC-----FEGDGSFLCGCRPGFLDDLVTC 300

RESULT 14
US-08-627-706-15
; Sequence 15, Application US/08627706
; Patent No. 5773696

```
; GENERAL INFORMATION:
; APPLICANT: Liang, Jihong
; APPLICANT: Shah, Dilip M.
; APPLICANT: Wu, Yinnie S.
; APPLICANT: Rosenberger, Cindy A.
; TITLE OF INVENTION: Antifungal Polypeptide and Methods for
; TITLE OF INVENTION: Controlling Plant Pathogenic Fungi
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F
; STREET: 700 Chesterfield Village Parkway No. 5773696th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,706
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Cohen, Charles E.
; REGISTRATION NUMBER: 34,565
; REFERENCE/DOCKET NUMBER: 38-21(10700)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6224
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-627-706-15
```

```
Query Match 17.7% Score 60; DB 1; Length 79;
Best Local Similarity 27.5%; Pred. No. 2.7;
Matches 19; Conservative 9; Mismatches 29; Indels 12; Gaps 4;

QY 4 IAIIFIVLVAFCILE-DGIVEAGF-----GCPFNAGKCHRHCKSIIR-RGGFCRGT 52
|:::| | | | | | | | | | | | | | | | | | | | |
Db 8 ISLLFAALVLFAPFAEPTMVDARLCERPSTGWSGCGNNNACRNLERAHSGSNV 67

QY 53 FRT-TCVCY 60
| | | | |
Db 68 FPAHKCICY 76
```

```
RESULT 15
US-09-103-489-15
; Sequence 15, Application US/09103489
; Patent No. 6215048
; GENERAL INFORMATION:
; APPLICANT: Liang, Jihong
; APPLICANT: Shah, Dilip M.
; APPLICANT: Wu, Yinnie S.
; APPLICANT: Rosenberger, Cindy A.
; TITLE OF INVENTION: Antifungal Polypeptide and Methods for
; TITLE OF INVENTION: Controlling Plant Pathogenic Fungi
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F
; STREET: 700 Chesterfield Village Parkway No. 6215048th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103,489
; FILING DATE: 24-JUN-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Cohen, Charles E.
; REGISTRATION NUMBER: 34,565
; REFERENCE/DOCKET NUMBER: 38-21 (10700)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 537-6224
; TELEFAX: (314) 537-6047
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-103-489-15
```

```
Query Match 17.7% Score 60; DB 4; Length 79;
Best Local Similarity 27.5%; Pred. No. 2.7;
Matches 19; Conservative 9; Mismatches 29; Indels 12; Gaps 4;

QY 4 IAIIFIVLVAFCILE-DGIVEAGF-----GCPFNAGKCHRHCKSIIR-RGGFCRGT 52
|:::| | | | | | | | | | | | | | | | | | | | |
Db 8 ISLLFAALVLFAPFAEPTMVDARLCERPSTGWSGCGNNNACRNLERAHSGSNV 67

QY 53 FRT-TCVCY 60
| | | | |
Db 68 FPAHKCICY 76
```

Search completed: September 17, 2002, 15:44:35
Job time: 126 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:09 ; Search time 29.91 seconds
(without alignments)
226.530 Million cell updates/sec

Title: US-09-829-481-4

Perfect score: 339

Sequence: 1 MKSIAIFIVLVAFCILEDG.....IRRRGFGCRGTFRTTCVCYR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

```

1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	177	52.2	38	19 AAW66437	Cationic peptide s
2	177	52.2	38	21 AAY91736	Cationic peptide s
3	133.5	39.4	38	15 AAR51160	Antibacterial pept
4	133.5	39.4	38	19 AAW66436	Cationic peptide i
5	133.5	39.4	38	21 AAY91735	Cationic peptide i
6	71.5	21.1	43	16 AAR82841	Allomycin A antiba
7	70	20.6	75	22 AAE10361	Soybean Gly m2 pro
8	70	20.6	94	10 AAP94261	Antibacterial poly
9	69.5	20.5	43	16 AAR82842	Allomycin B antiba
10	68.5	20.2	40	14 AAR30562	Sapecin - an antib
11	68.5	20.2	72	21 AAY44509	Pea Defensin prote

12	68	20.1	40	19 AAW66435	Cationic peptide s
13	68	20.1	40	21 AAY91734	Cationic peptide s
14	68	20.1	77	21 AAB24718	Plant SDF encoded
15	68	20.1	78	21 AAB24718	Plant SDF encoded
16	68	20.1	94	10 AAP90001	Antibacterial poly
17	68	20.1	100	21 AAB24717	Plant SDF encoded
18	67.5	19.9	43	18 AAW26000	Tenebrion. Tenebrio
19	67	19.8	567	21 AAB24967	Plant SDF encoded
20	67	19.8	799	21 AAB24966	Plant SDF encoded
21	67	19.8	817	21 AAB24965	Plant SDF encoded
22	66	19.5	40	11 AAR00720	Phormia terranova
23	66	19.5	40	19 AAW66454	Cationic peptide p
24	66	19.5	40	21 AAY91753	Cationic peptide p
25	66	19.5	45	15 AAR57129	Yeast pro-AMF C-te
26	65.5	19.3	84	21 AAG15344	Arabidopsis thalia
27	64.5	19.0	80	18 AAW19280	Raphanus sativus a
28	64.5	19.0	80	18 AAW19617	Radish antifungal
29	64.5	19.0	85	22 AAB20070	Scorpion sodium ch
30	63.5	18.7	54	22 AAU52922	Propionibacterium
31	63.5	18.7	80	22 AAB61792	Radish antifungal
32	63.5	18.7	81	20 AAY28848	Maize basal endosp
33	63.5	18.7	768	13 AAR27683	Rabbit beta-8 inte
34	63.5	18.7	768	13 AAR27684	Human beta-8 inte
35	63	18.6	88	21 AAG10630	Arabidopsis thalia
36	63	18.6	88	21 AAG45781	Arabidopsis thalia
37	63	18.6	90	21 AAG10629	Arabidopsis thalia
38	63	18.6	90	21 AAG45780	Arabidopsis thalia
39	63	18.6	106	21 AAG10628	Arabidopsis thalia
40	63	18.6	106	21 AAG45779	Arabidopsis thalia
41	62	18.3	27	21 AAY88036	Human growth hormo
42	61.5	18.1	77	21 AAG04820	Arabidopsis thalia
43	61.5	18.1	77	21 AAG38373	Arabidopsis thalia
44	61.5	18.1	77	21 AAY68682	Amino acid sequenc
45	61.5	18.1	80	14 AAR33706	Antifungal protein

ALIGNMENTS

RESULT 1
AAW66437
ID AAW66437 standard; peptide: 38 AA.

AC AAW66437;
DT 12-JAN-1999 (first entry)
XX
XX Cationic peptide scorpion defensin.
DE Indolicidin analogue; resistance; cationic peptide; antibiotic;
KW bacterial infection; tolerance; antibacterial; microorganism;
KW bacteria; fungus; parasite; virus.
XX
OS Lelurus quinquestriatus.
PN WO9840401-A2.
XX
PD 17-SEP-1998.
XX
PF 10-MAR-1998; 98WO-CA00190.
XX
PR 25-FEB-1998; 98US-0030619.
PR 10-MAR-1997; 97US-0040649.
PR 20-AUG-1997; 97US-0915314.
PR 26-SEP-1997; 97US-0060099.
XX
PA (MTCR-) MICROLOGIX BIOTECH INC.
XX
PI Fraser JR, McNicol PJ, West MHP;
XX WPI; 1998-520800/44.
XX
PT New Indolicidin peptide analogues - useful for, e.g. enhancing

PT activity of antibiotic or overcoming tolerance, acquired resistance
 XX or inherent resistance of microorganisms
 PS Disclosure; Page 10; 105pp; English.

XX AAW66393 to AAW66469 represent native cationic peptides from the
 CC present invention. The present invention describes compositions and
 CC methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses.

XX Sequence 38 AA;

Query Match 52.2%; Score 177; DB 19; Length 38;
 Best Local Similarity 73.0%; Pred. No. 1.9e-13;
 Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 25 GFGCPFNAGKCHRHCKSIIRRGFGCGRTTTCVCYR 61
 ||||| | |||||:||||||| | | | |||||
 Db 1 gfgcplnggachrhersirrrggycagffkqtctcyr 37

RESULT 2

AAAY91736
 ID AAY91736 standard; Peptide; 38 AA.

XX AAY91736;

XX 06-JUN-2000 (first entry)

DE Cationic peptide scorpion defensin amino acid sequence.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.

XX Unidentified.

XX WO9965506-A2.

XX 23-DEC-1999.

XX 14-JUN-1999; 99WO-CA00552.

XX 12-JUN-1998; 98US-0096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;

XX WPI; 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated
 PT polyoxyalkylene-modified cationic peptides, useful for treating tumours

XX Disclosure; Page 11; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically

CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.

XX Sequence 38 AA;

Query Match 52.2%; Score 177; DB 21; Length 38;
 Best Local Similarity 73.0%; Pred. No. 1.9e-13;
 Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 25 GFGCPFNAGKCHRHCKSIIRRGFGCGRTTTCVCYR 61
 ||||| | |||||:||||||| | | | |||||
 Db 1 gfgcplnggachrhersirrrggycagffkqtctcyr 37

RESULT 3

AAR51160
 ID AAR51160 standard; peptide; 38 AA.

XX AAR51160;

XX 18-OCT-1994 (first entry)

XX Antibacterial peptide induced in *Aeschna cyanea*.

DE Defensin; antibacterial peptide; Odonata; paleopteran insect;
 KW dragonfly.

XX *Aeschna cyanea*.

XX Key Location/Qualifiers

FT Disulfide-bond 4..26

FT Disulfide-bond 11..34

FT Disulfide-bond 14..36

FT Region 11..21

FT /label= alpha_helix

FT /note= "approximate position"

FT Region 24..38

FT /label= beta-pleated_sheet

FT /note= "approximate position"

XX FR2695392-A.

XX 11-MAR-1994.

XX 04-SEP-1992; 92FR-0010609.

XX 04-SEP-1992; 92FR-0010609.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Bulet P, Hetru C, Hoffmann J;

XX WPI; 1994-103535/13.

XX New antibacterial peptide(s) from dragonfly - for medical,
 PT veterinary, agricultural and food preservation use

XX Claim 10; Page 17; 25pp; French.

XX A defensin is induced in the dragonfly *Aeschna cyanea* by bacterial
 CC infection. The defensin (AAR51160) contains 3 disulphide bonds and has
 CC a structure comprising an alpha helix linked to a C-terminal
 CC antiparallel beta-pleated sheet through 2 of the disulphide bonds.
 CC The peptide has antibacterial activity which makes it suitable for
 CC use in the treatment of bacterial infections of eyes and ears and
 CC for oral/dental hygiene and gynaecology.

XX Sequence 38 AA;

Query Match 39.4%; Score 133.5; DB 15; Length 38;
 Best Local Similarity 55.3%; Pred. No. 1.9e-08;

Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFGCPNAGKCHRHCKSIR-RRGGFCRGTFRTTCVCYR 61
 ||||| : :|||::| | ||: | : || |||
 Db 1 gfgcpIdmqchrcqtigrsggycsgplklctcyr 38

RESULT 4
 AAW66436
 ID AAW66436 standard; peptide; 38 AA.
 XX
 AC AAW66436;
 XX
 DT 12-JAN-1999 (first entry)
 XX
 DE Cationic peptide insect defensin.
 XX
 KW Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KW bacterial infection; tolerance; antibacterial; microorganism;
 KW bacteria; fungus; parasite; virus.
 XX
 OS Aeschna cyanae.
 XX
 PN WO9840401-A2.
 XX
 PD 17-SEP-1998.
 XX
 PF 10-MAR-1998; 98WO-CA00190.
 XX
 PR 25-FEB-1998; 98US-0030619.
 PR 10-MAR-1997; 97US-0040649.
 PR 20-AUG-1997; 97US-0915314.
 PR 26-SEP-1997; 97US-0060099.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Fraser JR, McNicol PJ, West MHP;
 DR WPI; 1998-520800/44.
 XX
 PT New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX
 PS Disclosure; Page 10; 105pp; English.
 XX
 CC AAW66393 to AAW66469 represent native cationic peptides from the
 CC present invention. The present invention describes compositions and
 CC methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses.
 XX
 SQ Sequence 38 AA;

Query Match 39.4%; Score 133.5; DB 19; Length 38;
 Best Local Similarity 55.3%; Pred. No. 1.9e-08;
 Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFGCPNAGKCHRHCKSIR-RRGGFCRGTFRTTCVCYR 61
 ||||| : :|||::| | ||: | : || |||
 Db 1 gfgcpIdmqchrcqtigrsggycsgplklctcyr 38

RESULT 5
 AAY91735
 ID AAY91735 standard; Peptide; 38 AA.

XX AAY91735;
 AC
 XX 06-JUN-2000 (first entry)
 DT
 DE Cationic peptide insect defensin amino acid sequence.
 XX
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 XX
 OS Unidentified.
 XX
 PN WO9965506-A2.
 XX
 PD 23-DEC-1999.
 XX
 PF 14-JUN-1999; 99WO-CA00552.
 XX
 PR 12-JUN-1998; 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX
 DR WPI; 2000-223549/19.
 XX
 PR Novel pharmaceutical composition containing optionally activated
 PR polyoxyalkylene-modified cationic peptides, useful for treating tumours
 XX
 PS Disclosure; Page 11; 94pp; English.
 XX
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 XX
 SQ Sequence 38 AA;

Query Match 39.4%; Score 133.5; DB 21; Length 38;
 Best Local Similarity 55.3%; Pred. No. 1.9e-08;
 Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFGCPNAGKCHRHCKSIR-RRGGFCRGTFRTTCVCYR 61
 ||||| : :|||::| | ||: | : || |||
 Db 1 gfgcpIdmqchrcqtigrsggycsgplklctcyr 38

RESULT 6
 AAR82841
 ID AAR82841 standard; peptide; 43 AA.
 XX
 AC AAR82841;
 XX
 DT 25-MAR-1996 (first entry)
 XX
 DE Allomycin A antibacterial peptide.
 XX
 KW Allomycin; antibacterial; drug; food; recombinant.
 XX
 OS Allomyrina dichotoma.
 XX
 PN JP07196688-A.
 XX
 PD 01-AUG-1995.


```
XX 13-JAN-2000.
PD XX
XX 02-JUL-1999; 99WO-CA00608.
PF XX
XX 03-JUL-1998; 98CA-2242116.
PR XX
XX 06-JUL-1998; 98US-0091751.
XX XX
XX (UYMA-) UNIV MANITOBA.
XX XX
XX Fristensky B, Wang Y;
PI XX
XX WPI; 2000-126938/11.
DR XX
XX N-PSDB; AA229412.
XX XX
XX Recombinant expression system for expressing DRR206 or defensin, used
PT to produce pathogen resistant Brassica napus -
XX XX
XX Claim 11; Fig 9; 39pp; English.
XX XX
XX The present amino acid sequence is the defensin protein, derived from
CC the clone Drr230 of Pea. This sequence is induced by bacterial and
CC fungal pathogens like, blackleg fungus. The gene sequence is used in a
CC recombinant expression system, capable of transforming plants like,
CC Canola (Brassica napus), under the constitutive control of CaMV 35S
CC inducible promoter, responsive to pathogen infections. T-DNA sequence is
CC also present, for integration of the expression system into the plant
CC genome. Transgenic plants expressing DRR206 and defensin protein,
CC inhibited fungal growth in-vitro and are resistant to pathogenic
CC infections of Rhizoctonia solani, Leptosphaeria maculans and Sclerotinia
CC sclerotiorum.
XX XX
SQ Sequence 72 AA;

Query Match 20.2%; Score 68.5; DB 21; Length 72;
Best Local Similarity 28.4%; Pred. No. 1;
Matches 19; Conservative 11; Mismatches 24; Indels 13; Gaps 3;

QY 2 KSTAITIVLVAPCILEDGIVEAGFCGCPFNAGK-----CHRHCKSIIRRRG---GFCR 50
DB ||:|: |: |:| | | | | | | | | | | | | | | | | | | | | | | |
4 kslacIsflllvfvaqeivvseantcenlagsgkvcfggcdrhrt--qegaisgrcr 61

QY 51 GTFRTTC 57
DB || |
62 ddfrcwc 68

RESULT 12
AAW66435
ID AAW66435 standard; peptide; 40 AA.
XX XX
XX AAW66435;
AC XX
XX 12-JAN-1999 (first entry)
DT XX
XX Cationic peptide sapecin.
DE XX
XX Indolicidin analogue; resistance; cationic peptide; antibiotic;
KW bacterial infection; tolerance; antibacterial; microorganism;
KW bacteria; fungus; parasite; virus.
XX XX
XX Sacrophaga peregrina.
OS XX
XX WO9840401-A2.
PN XX
XX 17-SEP-1998.
PD XX
XX 10-MAR-1998; 98WO-CA00190.
PF XX
XX 25-FEB-1998; 98US-0030619.
PR XX
XX 10-MAR-1997; 97US-0040649.
PR 20-AUG-1997; 97US-0915314.
XX XX

26-SEP-1997; 97US-0060099.
XX (MICR-) MICROLOGIX BIOTECH INC.
XX XX
XX Fraser JR, McNicol PJ, West MHP;
PI XX
XX WPI; 1998-520800/44.
XX XX
XX New indolicidin peptide analogues - useful for, e.g. enhancing
PT activity of antibiotic or overcoming tolerance, acquired resistance
PT or inherent resistance of microorganisms
XX XX
XX Disclosure; Page 10; 105pp; English.
XX XX
XX AAW66393 to AAW66469 represent native cationic peptides from the
CC present invention. The present invention describes compositions and
CC methods for treating infection, especially bacterial infections. The
CC compositions and methods use cationic peptides in combination with an
CC antibiotic agent which are then administered to a patient to enhance the
CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
CC acquired resistance; and (c) inherent resistance. The combinations of
CC antibiotics and cationic peptides can provide synergistic activity
CC against a microorganism that is tolerant, inherently resistant, or has
CC acquired resistance to an antibiotic agent. They can be used for killing
CC e.g. bacteria, fungi, parasites and viruses.
XX XX
SQ Sequence 40 AA;

Query Match 20.1%; Score 68; DB 19; Length 40;
Best Local Similarity 39.4%; Pred. No. 0.65; Mismatches 16; Indels 2; Gaps 1;
Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;

QY 27 GCFPNAGKCHRHCKSIIRRRGGFCGRTTCVC 59
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
8 gTgInhsacaahellrgnrggycng--kavcvc 38

RESULT 13
AA91734
ID AAY91734 standard; Peptide; 40 AA.
XX XX
XX AAY91734;
AC XX
XX 06-JUN-2000 (first entry)
DT XX
XX Cationic peptide Sapecin amino acid sequence.
DE XX
XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma,
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.
XX XX
XX Unidentified.
OS XX
XX WO9965506-A2.
PN XX
XX 23-DEC-1999.
PD XX
XX 14-JUN-1999; 99WO-CA00552.
PF XX
XX 12-JUN-1998; 98US-0096541.
PR XX
XX (MICR-) MICROLOGIX BIOTECH INC.
PA XX
XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
PI XX
XX WPI; 2000-223549/19.
XX XX
XX Novel pharmaceutical composition containing optionally activated
PT polyoxyalkylene-modified cationic peptides, useful for treating tumours
PT -
XX XX
```

PS Disclosure; Page 11; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which

CC can be used in the pharmaceutical composition of the invention. The

CC invention relates to a pharmaceutical composition containing at least one

CC activated polyoxyalkylene (APO)-modified cationic peptide. The

CC modification of peptides with APO increases their activity against tumour

CC cells, including those with a multidrug resistant phenotype. The

CC pharmaceutical composition can be used to treat tumours, specifically

CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,

CC cervix, uterus, skin, prostate, liver and colon.

XX Sequence 40 AA;

Query Match 20.1%; Score 68; DB 21; Length 40;

Best Local Similarity 39.4%; Pred. No. 0.65; Mismatches 2; Gaps 1;

Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;

Oy 27 GCPFNAGKCHRHCKSIIRRGGFCRGTFRTTCVC 59

Db 8 gtginhsacaahcillrnnrggycng--kavcvc 38

RESULT 14

AAB24719

ID AAB24719 standard; Peptide; 77 AA.

AC AAB24719;

XX 27-NOV-2000 (first entry)

DT Plant SDF encoded polypeptide sequence SEQ List 1 NO:162.

DE Plant; corn; Arabidopsis thaliana; sequence-determined DNA fragment;

KW SDF; genetic mapping; identification; promoter; structural gene; UTR;

KW untranslated region; expression control.

XX OS

XX Plant.

PN WO2000040695-A2.

XX 13-JUL-2000.

XX 07-JAN-2000; 2000WO-US00466.

PF 08-JAN-1999; 99US-0115293.

XX (CERE-) CERES INC.

PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;

PI Zheng L;

XX WPI; 2000-465970/40.

XX New corn plant and Arabidopsis thaliana sequence-determined DNA

PT fragments, useful for expressing gene products and for controlling

PT expression of a target gene -

XX Claim 14; Page 400; 673pp; English.

XX The present invention describes polynucleotides, such as complete cDNA

CC sequences and/or sequences of genomic DNA encompassing complete genes,

CC portions of genes, and/or intergenic regions, collectively referred to

CC as sequence-determined DNA fragments (SDFs), from corn plants and

CC Arabidopsis thaliana. The SDFs are promoters, structural genes,

CC untranslated regions (UTRs), or 3' termination sequences. They can be

CC used for expressing a gene product and controlling expression of a

CC target gene, either as a promoter, a structural gene, an UTR or as a

CC 3' termination sequence. They are also useful as tools for genetic

CC mapping, and identification of a particular individual plant or for

CC clustering a group of plants with a common trait. AAA78433 to AAA78630

CC and AAB24605 to AAB25099 represent the specifically claimed

CC polynucleotide sequences and polypeptides encoded by them given in the

CC present invention.

SQ Sequence 77 AA;

Query Match 20.1%; Score 68; DB 21; Length 77;

Best Local Similarity 30.1%; Pred. No. 1.3; Mismatches 9; Gaps 4;

Matches 22; Conservative 9; Mismatches 26; Indels 16; Gaps 4;

Oy 2 KSTAIIFVLVAFCILEDGIV-EAGFGCCPFN-----AGKCHRHCKSIIRRG-G 47

Db 3 kllqvstvmiiftlilvlgvvaneglgpkkqcnellkqsnvcvaecdsmcvkkrkgag 62

Oy 48 FCRGTFRTTCVCY 60

Db 63 ycpspkk--cycy 73

RESULT 15

AAB24718

ID AAB24718 standard; Peptide; 78 AA.

XX AAB24718;

XX 27-NOV-2000 (first entry)

DT Plant SDF encoded polypeptide sequence SEQ List 1 NO:161.

DE Plant; corn; Arabidopsis thaliana; sequence-determined DNA fragment;

KW SDF; genetic mapping; identification; promoter; structural gene; UTR;

KW untranslated region; expression control.

XX OS

XX Plant.

PN WO2000040695-A2.

XX 13-JUL-2000.

XX 07-JAN-2000; 2000WO-US00466.

PF 08-JAN-1999; 99US-0115293.

XX (CERE-) CERES INC.

PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;

PI Zheng L;

XX WPI; 2000-465970/40.

XX New corn plant and Arabidopsis thaliana sequence-determined DNA

PT fragments, useful for expressing gene products and for controlling

PT expression of a target gene -

XX Claim 14; Page 399; 673pp; English.

XX The present invention describes polynucleotides, such as complete cDNA

CC sequences and/or sequences of genomic DNA encompassing complete genes,

CC portions of genes, and/or intergenic regions, collectively referred to

CC as sequence-determined DNA fragments (SDFs), from corn plants and

CC Arabidopsis thaliana. The SDFs are promoters, structural genes,

CC untranslated regions (UTRs), or 3' termination sequences. They can be

CC used for expressing a gene product and controlling expression of a

CC target gene, either as a promoter, a structural gene, an UTR or as a

CC 3' termination sequence. They are also useful as tools for genetic

CC mapping, and identification of a particular individual plant or for

CC clustering a group of plants with a common trait. AAA78433 to AAA78630

CC polynucleotide sequences and polypeptides encoded by them given in the

CC present invention.

XX Sequence 78 AA;

Query Match 20.1%; Score 68; DB 21; Length 78;
Best Local Similarity 30.1%; Pred. NO. 1.3;
Matches 22; Conservative 9; Mismatches 26; Indels 16; Gaps 4;
QY 2 KSIAIIFIVLVAFCILEGIV-EAGFGCPFN-----AGKCHRHCKSIRRG-G 47
DB 4 kllqvsftvmiiftllvgvvaneglgkpkqkneilkqsnvcvaaecdsmcvkkrkgag 63
QY 48 FCRGTFRRTTCVCY 60
DB 64 ycspskk--cycy 74

Search completed: September 17, 2002, 15:44:15
Job time: 126 sec

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	184	54.3	37	1	DEF4_ANDAU	P56686 androctonus
2	177	52.2	38	1	DEF4_LEIOH	P41965 leirurus qui
3	133.5	39.4	38	1	DEF1_AESCY	P80154 aeschna cya
4	128	37.8	38	1	DEF1_MYGA	P80571 mytilus gal
5	127	37.5	37	1	DEF1_MYTED	P81610 mytilus edu
6	114	33.6	35	1	DEFB_MYTED	P81611 mytilus edu
7	82.5	24.3	85	1	SC11_MESMA	Q9njc7 mesobuthus
8	80.5	23.7	43	1	DEF4_ZOPAT	P80033 zophobas at
9	79.5	23.5	85	1	SCX2_ANDAU	P01484 androctonus
10	75.5	22.3	85	1	SCAL_MESMA	O9gyx2 mesobuthus
11	71.5	21.1	43	1	DEF1_ALLDI	O10745 allomyrina
12	71.5	21.1	74	1	DEF2_CAPAN	O65740 capsicum an
13	71.5	21.1	84	1	DEF1_TENMO	O27023 tenebrio mo
14	71	20.9	99	1	DEF3_AEDAE	P81603 aedes aegypt
15	70.5	20.8	78	1	THGF_HELAN	P82659 helianthus
16	69.5	20.5	65	1	SCYB_BUTOC	P01486 buthus occi
17	69	20.4	75	1	LOKD_VIGUN	P18646 vigna ungui
18	69	20.4	102	1	DEF1_ANOGA	Q17027 anopheles g
19	68.5	20.2	72	1	D230_PEA	P01783 pisum sativ
20	68.5	20.2	85	1	SC15_MESMA	Q9qng8 mesobuthus
21	68	20.1	94	1	SAPF_SARPE	P18313 sarcophaga
22	66.5	19.6	64	1	SCX3_BUTOC	P01485 buthus occi
23	66.5	19.6	86	1	SCX7_CENNO	P45665 centruroides
24	66	19.5	94	1	DEF1_PROTE	P10891 protophormi
25	65.5	19.3	64	1	SCX5_ANDMA	P01482 androctonus
26	65.5	19.3	64	1	SCX5_LEIOU	P01481 leirurus qui
27	64.5	19.0	798	1	ITB7_HUMAN	P26010 homo sapien
28	63.5	18.7	77	1	THG2_ARATH	Q41914 arabidopsis
29	63.5	18.7	80	1	AFP2_RAPSA	P30230 raphanus sa
30	63.5	18.7	94	1	SCR9_PANIM	P56972 pandinus im
31	63.5	18.7	768	1	ITB8_RABIT	P26013 cryctolagus
32	63.5	18.7	769	1	ITB8_HUMAN	P26012 homo sapien
33	63	18.6	40	1	DEFB_AEDAE	P81602 aedes aegypt

OC Buttho

```

OX NCBI_TaxID=6884;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=93326112; PubMed=8333834;
RA Cociancich S., Goffion M., Bontems F., Bulet P., Bouet F., Menez A.,
  Hoffmann J.A.;
RT "Purification and characterization of a scorpion defensin, a 4kDa
  antibacterial peptide presenting structural similarities with insect
  defensins and scorpion toxins.";
RL Biochem. Biophys. Res. Commun. 194:17-22(1993).
CC -!- FUNCTION: ANTIBACTERIAL PROTEIN AGAINST GRAM-POSITIVE BACTERIA;
CC MAY ACT VIA MEMBRANE-PERMEABILISATION OF THESE CELLS.
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR PIR; JN0613; JN0613.
DR HSP; P10891; 1ICA.
DR InterPro; IPR001542; Arthro_defensin.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Antibiotic.
FT DISULFID 4 25 BY SIMILARITY.
FT DISULFID 11 33 BY SIMILARITY.
FT DISULFID 15 35 BY SIMILARITY.
SQ SEQUENCE 38 AA; 4326 MW; DF35FB21ECB3FB84 CRC64;

Query Match 52.2%; Score 177; DB 1; Length 38;
Best Local Similarity 73.0%; Pred. No. 1.4e-13;
Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

-QY 25 GFGCPFNAGKCHRRCKSIRRRGFCGRTFTTCVCYR 61
  ||||| | | ||||| | | ||||| | | |||||
Db 1 GFGCPNAGKCHRRCKSIRRRGFCGRTFTTCVCYR 37

-RESULT 3
ID DEFLAESC STANDARD; PRT; 38 AA.
AC P80154;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Defensin.
OS Aeschna cyanea (Dragonfly) (Blue darner).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Palaeoptera; Odonata; Anisoptera; Aeshnidae; Aeshna.
OX NCBI_TaxID=12921;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=93049356; PubMed=1425705;
RA Bulet P., Cociancich S., Reuland M., Sauber F., Bischoff R., Hegy G.,
  van Dorsseleer A., Hetru C., Hoffmann J.A.;
RT "A novel insect defensin mediates the inducible antibacterial
  activity in larvae of the dragonfly Aeschna cyanea (Paleoptera,
  Odonata).";
RL Eur. J. Biochem. 209:977-984(1992).
CC -!- FUNCTION: MEDIATES THE INDUCIBLE ANTIBACTERIAL ACTIVITY IN LARVAE
  OF A. CYANEA.
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR PIR; S27242; S27242.
DR InterPro; IPR001542; Arthro_defensin.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Insect immunity; Antibiotic.
FT DISULFID 4 26 BY SIMILARITY.
FT DISULFID 11 34 BY SIMILARITY.
FT DISULFID 15 36 BY SIMILARITY.
SQ SEQUENCE 38 AA; 4180 MW; 7C884A8DF54A46A CRC64;

Query Match 39.4%; Score 133.5; DB 1; Length 38;
Best Local Similarity 55.3%; Pred. No. 9.7e-09;

```

```

Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFGCPFNAGKCHRRCKSIRRRGFCGRTFTTCVCYR 61
  ||||| | | ||||| | | ||||| | | |||||
Db 1 GFGCPNAGKCHRRCKSIRRRGFCGRTFTTCVCYR 38

RESULT 4
ID DEFLMYTGA STANDARD; PRT; 38 AA.
AC P80571;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Defensin MGD-1.
GN PH3.
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=97025339; PubMed=8925841;
RA Hubert F., Noeel T., Roch P.;
RT "A member of the arthropod defensin family from edible Mediterranean
  mussels (Mytilus galloprovincialis).";
RL Eur. J. Biochem. 240:302-306(1996).
RN [2]
RP ERRATUM.
RA Hubert F., Noeel T., Roch P.;
RL Eur. J. Biochem. 240:815-815(1996).
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR InterPro; IPR001542; Arthro_defensin.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Antibiotic.
FT DISULFID 4 25 BY SIMILARITY.
FT DISULFID 10 33 BY SIMILARITY.
FT DISULFID 14 35 BY SIMILARITY.
FT DISULFID 21 38 BY SIMILARITY.
SQ SEQUENCE 38 AA; 4275 MW; 6F1C0CDCF0E69E76 CRC64;

Query Match 37.8%; Score 128; DB 1; Length 38;
Best Local Similarity 63.2%; Pred. No. 4e-08;
Matches 24; Conservative 2; Mismatches 10; Indels 2; Gaps 2;

QY 25 GFGCPFNAGKCHRRCKSIRRRGFCGRTFTTCVCYR 61
  ||||| | | ||||| | | ||||| | | |||||
Db 1 GFGCPNAGKCHRRCKSIRRRGFCGRTFTTCVCYR 37

RESULT 5
ID DEFLMYTED STANDARD; PRT; 37 AA.
AC P81610;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Defensin A.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE AND CHARACTERIZATION.
RC TISSUE=Blood;
RX MEDLINE=96355569; PubMed=8702979;
RA Charlet M., Chernysh S., Philippe H., Hetru C., Hoffmann J.A.,
  Bulet P.;
RT "Innate immunity. Isolation of several cysteine-rich antimicrobial
  peptides from the blood of a mollusc, Mytilus edulis.";

```

RESULT 7
SC11_MESMA

```

RESULT      8
DEFA_ZOPAT
ID_DEFA_ZOPAT
AC P80033;
STANDARD;
PRT;      43 AA.

```

```

DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Defensin, isoforms B and C.
OS Zophobas atratus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OX Cucujiformia; Tenebrionidae; Zophobas.
OX NCBI_TaxID=7074;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=92105112; PubMed=1761552;
RA Bulet P., Cocianich S., Dimarcq J.-L., Lambert J., Reichhart J.-M.,
RA Hofmann D., Hetré C., Hoffmann J.A.;
RT "Insect immunity. Isolation from a coleopteran insect of a novel
RT inducible antibacterial peptide and of new members of the insect
RT defensin family."
RL J. Biol. Chem. 266:24520-24525(1991).
CC -1- FUNCTION: INVOLVED IN ANTI GRAM-POSITIVE ACTIVITY OF IMMUNE
CC -1- HEMOLYMPH OF Z.ATRATUS.
CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR PIR; B41711; B41711.
DR PIR; C41711; C41711.
DR HSSP; P10891; LICA.
DR InterPro; IPR001542; Arthro_defensin.
DR InterPro; IPR003614; Knott1.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PRINTS; PR00271; DEFENSIN.
DR SMART; SM00505; Knott1; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Insect immunity; Antibiotic.
FT DISULFID 3 34 BY SIMILARITY.
FT DISULFID 20 40 BY SIMILARITY.
FT DISULFID 24 42 BY SIMILARITY.
FT VARIANT 30 30 R -> T (IN ISOFORM C).
SQ SEQUENCE 43 AA; 4453 MW; D63E7B100987FD2F CRC64;

Query Match 23.7%; Score 80.5; DB 1; Length 43;
Best Local Similarity 39.5%; Pred. No. 0.0088;
Matches 15; Conservative 6; Mismatches 12; Indels 5; Gaps 2;

QY 25 GF---GPFNAGKCHRRCKSIRRGSGCFRTFTTCVC 59
|| | | | | | | | | | | | | | | | | | |
Db 7 GFETAGTKLSAAGAHCHALGRGGYCNS--KSVCCVC 42

RESULT 9
SCX2_ANDAU STANDARD; PRT; 85 AA.
AC P01484;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurotoxin II precursor (Aah II).
OS Androctonus australis hector (Sahara scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butchoidea; Butidae; Androctonus.
OX NCBI_TaxID=70175;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90037062; PubMed=2808423;
RA Bougis P.E., Rochat H., Smith L.A.;
RT "Precursors of Androctonus australis scorpion neurotoxins. Structures
RT of precursors, processing outcomes, and expression of a functional
RT recombinant toxin II."
RL J. Biol. Chem. 264:19259-19265(1989).
RN [2]
RP SEQUENCE OF 20-82.
RX MEDLINE=73025153; PubMed=4342910;
RA Rochat H., Rochat C., Sampieri F., Miranda F., Lissitzky S.;
RT "The amino-acid sequence of neurotoxin II of Androctonus australis

```

```

hector."
RT Eur. J. Biochem. 28:381-388(1972).
RN [3]
RP DISULFIDE BONDS.
RX MEDLINE=75057756; PubMed=4611766;
RA Kopeyan C., Martinez G., Lissitzky S., Miranda F., Rochat H.;
RT "Disulfide bonds of toxin II of the scorpion Androctonus australis
RT hector."
RL Eur. J. Biochem. 47:483-489(1974).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS).
RX MEDLINE=94194522; PubMed=8145259;
RA Housset D., Habersetzer-Rochat C., Astier J.P., Fontecilla-Camps J.C.;
RT "Crystal structure of toxin II from the scorpion Androctonus
RT australis hector refined at 1.3-A resolution."
RL J. Mol. Biol. 238:88-103(1994).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (0.96 ANGSTROMS).
RA Smith G.D., Blessing R.H., Ealick S.E., Fontecilla-Camps J.C.,
RA Hauptman H.A., Housset D., Lango D.A., Miller R.;
RT "Ab initio structure determination and refinement of a scorpion
RT protein toxin."
RL Acta Crystallogr. D 53:551-557(1997).
CC -1- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBITS THE INACTIVATION
CC OF THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M27704; AAA29949.1; -
DR PIR; A01744; NTSR2A.
DR PIR; D34444; D34444.
DR PDB; 1PTX; 26-JAN-95.
DR PDB; 1AHO; 15-OCT-97.
DR InterPro; IPR003614; Knott1.
DR InterPro; IPR002061; Scorpion_toxin.
DR InterPro; IPR001219; Toxin.
DR Pfam; PF00537; toxin_3; 1.
DR PRINTS; PR00284; TOXIN.
DR PRODOM; PD000908; Scorpion_toxin; 1.
DR SMART; SM00505; Knott1; 1.
KW Neurotoxin; Amidation; Sodium channel inhibitor; Signal; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 83 NEUROTOXIN II.
FT DISULFID 31 82
FT DISULFID 35 55
FT DISULFID 41 65
FT DISULFID 45 67
FT MOD_RES 83 83
SQ SEQUENCE 85 AA; 9548 MW; 1740CC6B98363768 CRC64;

Query Match 23.5%; Score 79.5; DB 1; Length 85;
Best Local Similarity 24.2%; Pred. No. 0.019;
Matches 16; Conservative 16; Mismatches 27; Indels 7; Gaps 3;

QY 1 MKSAITFIIVLVAFCTLEDGIVEAGFCGFNAGK---CHRHCKSIRRGSGCFR--GTERT 55
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6 MISALLFVTGVE--SVKDGIVDDVNCYFCGRNAYCNFECKLKGSGYQWASPIGN 63

QY 56 TCVCYR 61
| | |
Db 64 ACYCYK 69

```

OS	Allomyrina dichotoma.
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC	Scarabaeiformia; Scarabaeidae; Dynastinae; Allomyrina.
OX	NCBI_Taxid=50413;
[1]	
RN	SEQUENCE.
RP	TISSUE=Hemolymph;
RC	MEDLINE=96182985; PubMed=8607799;
RA	Miyashita A., Hara S., Sugiyama M., Asaoka A., Taniiai K.,
RY	Yokuburo F., Yamakawa M.;
RT	"Isolation and characterization of a new member of the insect
RT	defensin family from a beetle, <i>Allomyrina dichotoma</i> .";
RL	Biochem. Biophys. Res. Commun. 220:526-531(1996).
CC	-!- FUNCTION: ANTIBACTERIAL PEPTIDE. STAPHYLOCOCCUS BACTERIA
CC	INCLUDING METHICILLIN-RESISTANT STAPHYLOCOCCUS AUREUS.
CC	-!- INDUCTION: BY BACTERIAL INFECTION.
CC	-!- MASS SPECTROMETRY: MW=4542; MWERR=1.4; METHOD=Electrospray.
CC	-!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR	HSPSP; P10891; LICA.
DR	InterPro: IPR001542; Arthro_defensin.
DR	InterPro: IPR003614; KnottI.
DR	Pfam: PF01097; Arthro_defensin; 1.
DR	PRINTS; PR00271; DEFENSIN.
DR	SMART; SW00505; KnottI; 1
DR	PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW	Insect Immunity; Antibiotic.
FT	DISULFID 3 34 BY SIMILARITY.
FT	DISULFID 20 39 BY SIMILARITY.
FT	DISULFID 24 41 BY SIMILARITY.
SQ	SEQUENCE 43 AA; 4553 MW; 5407246E9E673941 CRC64;

Query Match	21.1%;	Score 71.5;	DB 1;	Length 43;
Best Local Similarity	42.9%;	Pred. No. 0.089;		
Matches 15;	Conservative	2; Mismatches	15;	Indels 3; Gap

Qy	27	GCPFNACKCHRHCKSIIRRGSGCRGTFRTTCVCYCR 61
	:	
Dd	12	GFAANHSLCAAHCLAIARRGGSCSE---RGCVICRR 43

RESULT 12	
DEF2_CAPAN	
ID DEF2_CAPAN	STANDARD; PRT; 74 AA.
AC O65740;	
DT 16-OCT-2001 (Rel. 40, Created)	
DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DT 01-MAR-2002 (Rel. 41, Last annotation update)	
DE Defensin J1-2 precursor.	
OS Capsicum annuum (Bell pepper).	
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;	
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.	
ON NCBI_TaxID=4072;	
XX [1]	
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.	
RC STRAIN=CV_YOLO WONDER; TISSUE=Fruit;	
RX MEDLINE=97037730; PubMed=8883377;	
RA Meyer B., Houline G., Pozueta-Romero J., Schantz M.L., Schantz R.;	
RT "Fruit-specific expression of a defensin-type gene family in bell	
RT pepper. Upregulation during ripening and upon wounding.";	
RL Plant Physiol. 112:615-622(1996).	
[2]	
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.	
RC STRAIN=CV_YOLO WONDER; TISSUE=Fruit;	
RX MEDLINE=99005242; PubMed=9790581;	
RA Houline G., Meyer B., Schantz R.;	
RT "Alteration of the expression of a plant defensin gene by exon	
RT shuffling in bell pepper (<i>Capsicum annum</i> L.).";	
RL Mol. Gen. Genet. 259:504-510(1998).	
CC -!- FUNCTION: PLANT DEFENSE PEPTIDE WITH ANTI-FUNGAL ACTIVITY AGAIN	
CC F.OXYSPORUM AND B.CINEREA.	

```

CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN FLOWERS AND IN YOUNG FRUITS.
CC -!- DEVELOPMENTAL STAGE: ACCUMULATE DURING RIPENING.
CC -!- SIMILARITY: BELONGS TO THE GAMMA-PUROTHIONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95730; CAA65046.1; -.
DR HSSP; P20230; IGPT.
DR Mendel; 29862; Capan;1533;29862.
DR InterPro; IPR002118; Gamma-thionin.
DR Pfam; PR003614; Knot1.
DR ProDom; PD002594; Gamma-thionin; 1.
DR SMART; SM00505; Knot1; 1.
DR PROSITE; PS00940; GAMMA_THIONIN; 1.
KW Defensin; Fungicide; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 74 DEFENSIN J1-2.
FT DISULFID 30 74 BY SIMILARITY.
FT DISULFID 41 61 BY SIMILARITY.
FT DISULFID 47 68 BY SIMILARITY.
FT DISULFID 51 70 BY SIMILARITY.
SQ SEQUENCE 74 AA; 8249 MW; D92D8F06F39E1552 CRC64;

Query Match 21.1%; Score 71.5; DB 1; Length 74;
Best Local Similarity 30.0%; Pred. No. 0.13;
Matches 21; Conservative 8; Mismatches 28; Indels 13; Gaps 3;

QY 2 KSAIIFVLVAFCLIEDGIVEAGFGCP-----FNAGKCHRHCKSIIRRGFGCRG 51
Db ||| |||::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
6 KVIATIFLMM--LVPATGMVAEATCESQSHRFKGLCFKSCNCGSVCHTEGFGNGHCRG 63
QY 52 TPTTCVCYR 61
Db || | | |
64 -FRRRCFCYR 72

RESULT 13
DEFL_TENMO
ID DEFL_TENMO STANDARD; PRT; 84 AA.
AC Q27023;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Tenebrion 1 precursor.
OS Tenebrio molitor (yellow mealworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Tenebrionidae; Tenebrio.
OX NCBI_TaxID=7067;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 49-59, AND DISULFIDE BONDS.
RC TISSUE=Larval hemolymph;
RX MEDLINE=95096025; PubMed=7798186;
RA Moon H.J., Lee S.Y., Kurata S., Natori S., Lee B.L.;
RT "Purification and molecular cloning of cDNA for an inducible
RT antibacterial protein from larvae of the coleopteran, Tenebrio
RT molitor.";
RL J. Biochem. 116:53-58(1994).
CC -!- FUNCTION: BACTERICIDAL PROTEIN PRODUCED IN RESPONSE TO INJURY. IT
CC IS CYTOTOXIC TO GRAM-POSITIVE BACTERIA.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17670; BAA04552.1; -.
DR HSSP; P10891; LICA.
DR InterPro; IPR001542; Arthro_defensin.
DR InterPro; IPR003614; Knot1.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PRINTS; PR00271; DEFENSIN.
DR SMART; SM00505; Knot1; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Insect immunity; Antibiotic; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 41
FT CHAIN 42 84 TENECIN 1.
FT DISULFID 44 75
FT DISULFID 61 81
FT DISULFID 65 83
SQ SEQUENCE 84 AA; 9176 MW; 0367C5070468BE60 CRC64;

Query Match 21.1%; Score 71.5; DB 1; Length 84;
Best Local Similarity 27.3%; Pred. No. 0.15;
Matches 21; Conservative 7; Mismatches 26; Indels 23; Gaps 3;

QY 4 IAIIFVLVAFCLIEDGI---VEAG-----FGCPFNAGKCHRHCKSI 42
Db ||| |||::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
9 VACPFILQIAAFPLEEAATAETEGSEHVRKRVCTDILSVEAKGKLNDAACAAHCLFR 68
QY 43 RRRGGFCRGTFRTVCV 59
Db ||| |||::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
69 GRGGYCNG--KRVVCV 83

RESULT 14
DEFL_AEDAE
ID DEFL_AEDAE STANDARD; PRT; 99 AA.
AC P81603; Q9Y0F1; Q9Y0F0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Defensin C precursor.
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, INDUCTION, AND SUBCELLULAR
RP LOCATION.
RC STRAIN=LIVERPOOL; TISSUE=Fat body;
RX MEDLINE=99124369; PubMed=9927179;
RA Lowenberger C.A., Smart C.T., Bulet P., Ferdig M.T., Severson D.W.,
RA Hoffmann J.A., Christensen B.M.;
RT "Insect immunity: molecular cloning, expression, and characterization
RT of cDNAs and genomic DNA encoding three isoforms of insect defensin in
RT Aedes aegypti.";
RL Insect Mol. Biol. 8:107-118(1999).
RN [2]
RP SEQUENCE OF 60-99, AND INDUCTION.
RC STRAIN=LIVERPOOL;
RX MEDLINE=95360030; PubMed=7633471;
RA Lowenberger C., Bulet P., Charlet M., Hetru C., Hodgeman B.,
RA Christensen B.M., Hoffmann J.A.;
RT "Insect immunity: isolation of three novel inducible antibacterial
RT defensins from the vector mosquito, Aedes aegypti.";
RL Insect Biochem. Mol. Biol. 25:867-873(1995).
RN [3]

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:49 ; Search time 18.35 seconds
(without alignments)
319.425 Million cell updates/sec

Title: US-09-829-481-4
Perfect score: 339
Sequence: 1 MKSAIIFIVLVAFCILEDG.....IRRGFGCGTFTTCVCYR 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	177	52.2	38	2 JN0613	defensin 4K - scor
2	133.5	39.4	38	2 S27242	defensin - blue da
3	128	37.8	38	2 S74088	defensin - Mediter
4	80.5	23.7	43	2 B41711	defensin B - beetl
5	79.5	23.5	85	1 NTSR2A	neurotoxin II prec
6	74.5	22.0	43	2 C41711	defensin C - beetl
7	72.5	21.4	43	2 JC2554	holotricin I - Hol
8	71.5	21.1	84	2 JX0332	tenecin I precursor
9	69.5	20.5	65	1 NTSREB	neurotoxin XI - sc
10	69	20.4	75	2 S11156	PSA10 protein - c
11	68.5	20.2	72	2 T06599	disease resistance
12	68	20.1	94	2 A31792	sapacin precursor
13	67.5	19.9	79	2 T06381	proteinase inhibit
14	66.5	19.6	64	1 NTSR3B	neurotoxin III - s
15	66.5	19.6	86	2 JN0671	Nat-channel-blocki
16	66	19.5	94	2 S12558	defensin A precurs
17	65.5	19.3	64	1 NTSR5L	neurotoxin V - Egy
18	65.5	19.3	64	1 NTSR5M	neurotoxin V - sco
19	64.5	19.0	798	2 A40526	integrin beta-7 ch
20	63.5	18.7	77	2 B84433	proteinase inhibit
21	63.5	18.7	80	2 T10823	antifungal protein
22	63.5	18.7	768	2 B41029	integrin beta-8 ch
23	63.5	18.7	769	2 A41029	integrin beta-8 ch
24	63	18.6	90	2 B84867	probable trypsin i
25	62	18.3	990	2 I51618	nucleolar phosphop
26	61.5	18.1	77	2 S30578	protease inhibitor
27	61.5	18.1	80	2 T10176	antifungal protein
28	61.5	18.1	154	2 T17816	hypothetical prote
29	61	18.0	40	2 JU0225	sapacin C - flesh

30 60.5 17.8 80 2 F96787 protein T4O12.7 [i
31 60.5 17.8 126 2 T21762 hypothetical prote
32 60.5 17.8 369 2 C75378 glutamate racemase
33 60.5 17.8 806 2 A46271 integrin beta-7 ch
34 60 17.7 79 2 T10243 antifungal protein
35 60 17.7 84 2 S62868 toxin gamma precu
36 59 17.4 84 2 S21158 neurotoxin TsVII p
37 58.5 17.3 75 2 S51637 probable proteinas
38 58.5 17.3 77 2 T14395 proteinase inhibit
39 58.5 17.3 80 2 T10183 antifungal protein
40 58.5 17.3 80 2 F96591 probable antifunga
41 58.5 17.3 640 2 T10243 protein kinase-lik
42 58 17.1 304 2 F69230 pyruvate formate-1
43 58 17.1 318 2 E87929 protein T22H2.6 [i
44 58 17.1 358 2 T25137 hypothetical prote
45 58 17.1 964 2 JC5545 integrin beta-4 pr

ALIGNMENTS

RESULT 1

JN0613

defensin 4K - scorpion (Leiurus quinquestriatus)

N:Alternate names: antibacterial 4K peptide

C:Species: Leiurus quinquestriatus hebraeus

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 03-Feb-1994

C:Accession: JN0613

R:Cociancich, S.; Goyffon, M.; Bontems, F.; Bulet, P.; Bouet, F.; Menez, A.; Hoffmann

Biochem. Biophys. Res. Commun. 194, 17-22, 1993

A:Title: Purification and characterization of a scorpion defensin, a 4kDa antibacteri

A:Reference number: JN0613; MUID:93326112

A:Accession: JN0613

A:Molecule type: protein

A:Residues: 1-38 <COC>

A>Note: This protein is similar to scorpion toxins and insect defensins

Query Match 52.2%; Score 177; DB 2; Length 38;

Best Local Similarity 73.0%; Pred. No. 5.4e-13;

Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 25 GFQCPNAGKCHRHCKSIRRRGGFCRGTFRTTCVCYR 61

Db 1 GFQCPNAGKCHRHCKSIRRRGGYCGAGFKQCTCYR 37

RESULT 2

S27242

defensin - blue darter

C:Species: Aeschna cyanea (blue darter)

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-1999

C:Accession: S27242

R:Bulet, P.; Cociancich, S.; Reuland, M.; Sauber, F.; Bischoff, R.; Hegy, G.; van Dor

Eur. J. Biochem. 209, 977-984, 1992

A:Title: A novel insect defensin mediates the inducible antibacterial activity in lar

A:Reference number: S27242; MUID:93049356

A:Accession: S27242

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-38 <BUL>

Query Match 39.4%; Score 133.5; DB 2; Length 38;

Best Local Similarity 55.3%; Pred. No. 3.6e-08;

Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFQCPNAGKCHRHCKSIR-RRGGFCRGTFRTTCVCYR 61

Db 1 GFQCPNAGKCHRHCKSIR-RRGGYCGAGFKQCTCYR 38

RESULT 3

A:Molecule type: protein
A:Residues: 20-83 <ROC>
R:Kopeyvan, C.; Martinez, G.; Lissitzky, S.; Miranda, F.; Rochat, H.
Eur. J. Biochem. 47, 483-489, 1974
A:Title: Disulfide bonds of toxin II of the scorpion Androctonus australis Hector.
A:Reference number: A91225; MUID:75057756
A:Contents: annotation; disulfide bonds
R:Housset, D.; Habersetzer-Rochat, C.; Astier, J.P.; Pontecilla-Camps, J.C.
J. Mol. Biol. 238, 88-103, 1994
A:Title: Crystal structure of toxin II from the scorpion Androctonus australis Hector
A:Reference number: A58108; MUID:94194522
A:Contents: annotation; X-ray crystallography, 1.3 angstroms, residues 20-83
R:Pontecilla-Camps, J.C.; Housset, D.
submitted to the Brookhaven Protein Data Bank, September 1994
A:Reference number: A67142; PDB:IPTX
A:Contents: annotation; X-ray crystallography, 1.3 angstroms, residues 20-83
C:Superfamily: scorpion neurotoxin
C:Keywords: amidated carboxyl end; neurotoxin; venom
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-83/Product: neurotoxin II #status experimental <MAT>
F:31-82;35-55;41-65;45-67/Disulfide bonds: #status experimental
F:83/Modified site: amidated carboxyl end (His) (amide in mature form from following

Query Match 23.5%; Score 79.5; DB 1; Length 85;
Best Local Similarity 24.2%; Pred. No. 0.062;
Matches 16; Conservative 16; Mismatches 27; Indels 7; Gaps 3;

QY 1 MKSIAIPIVLVAFCIILEDGIVEAGFCPPNAGK--CHRHCKSIIRRRGGFCR--GTFT 55
| | : : : : : | : : : : :
Db 6 MISLALLFTGVE--SVKDGVIYDVNVCTPCGRNAYCNBECKLKGESGYCWASPYGN 63

QY 56 TCVCYR 61
| | : : : : :
Db 64 ACYCIR 69

RESULT 6
C41711
defensin C - beetle (Zophobas atratus)
C:Species: Zophobas atratus
C:Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 03-May-1996
C:Accession: C41711
R:Bulet, P.; Cocciandich, S.; Dimarcq, J.L.; Lambert, J.; Reichhart, J.M.; Hoffmann, D.
J. Biol. Chem. 266, 24520-24525, 1991
A:Title: Insect immunity. Isolation from a coleopteran insect of a novel inducible an
A:Reference number: A41711; MUID:92105112
A:Accession: C41711
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-43 <BUI>
C:Superfamily: sapecin precursor

Query Match 22.0%; Score 74.5; DB 2; Length 43;
Best Local Similarity 36.8%; Pred. No. 0.14;
Matches 14; Conservative 6; Mismatches 13; Indels 5; Gaps 2;

QY 25 GF---GCCPFNAGKHRRHCKSIIRRRGGFCRGTFRTTCVC 59
| | : : : : : | : : : : :
Db 7 GFETAGTKLSNACGAHCLALGRGTGGYCNS--KSVCVC 42

RESULT 7
JC2554
holotricin 1 - Holotrichia diomphalia
C:Species: Holotrichia diomphalia
C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 17-Mar-1999
C:Accession: JC2554
R:Lee, S.Y.; Moon, H.J.; Kawabata, S.; Kurata, S.; Natori, S.; Lee, B.I.
Biol. Pharm. Bull. 18, 457-459, 1995
A:Title: A sapecin homologue of Holotrichia diomphalia: Purification, sequencing and
A:Reference number: JC2554; MUID:96031330

A:Accession: JC2554

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-43 <LEE>

C:Superfamily: sapecin precursor

Query Match 21.4%; Score 72.5; DB 2; Length 43;

Best Local Similarity 39.4%; Pred. No. 0.23; Mismatches 5; Gaps 1;

QY 27 GCPFNAGKCHRHCKSIIRRRGGFCGRTTCVC 59

Db 12 GIAINDSACAHCCLAMRRKGSGCK---QGVVC 41

RESULT 8

JX0332

tenecin 1 precursor - yellow mealworm

C:Species: Tenebrio molitor (yellow mealworm)

C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 20-Jun-2000

C:Accession: JX0332

R:Moore, H.J.; Lee, S.Y.; Kurata, S.; Natori, S.; Lee, B.L.

J. Biochem. 116, 53-58, 1994

A:Title: Purification and molecular cloning of cDNA for an inducible antibacterial prote

A:Reference number: JX0332; MUID:95096025

A:Accession: JX0332

A:Molecule type: mRNA

A:Residues: 1-84 <MOO>

A:Cross-references: GB:D17670; NID:g1235940; PIDN:BA04552.1; PID:g1235941

A:Experimental source: larva

C:Comment: Tenecin 1 is an antibacterial protein induced in the hemolymph of larvae of t

C:Superfamily: sapecin precursor

F:1-41/Domain: signal sequence #status predicted <SIG>

F:42-84/Product: tenecin 1 #status predicted <MAT>

F:44-75,61-81,65-83/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 21.1%; Score 71.5; DB 2; Length 84;

Matches 21; Conservative 7; Mismatches 26; Indels 23; Gaps 3;

QY 4 IAIIFIVLVAFPCILEDI---VEAG-----FGCPFNAGKCHRHCKSI 42

Db 9 VACFFILOIAFPLEEAATAAEIEQGEHIRVKRVTCDILSVEAKGVKLNDAACAAHCLFR 68

QY 43 RRRGGFCGRTTCVC 59

Db 69 GRSGGYCNG--KRVVC 83

RESULT 9

NTSREB

neurotoxin XI - scorpion (Buthus occitanus)

C:Species: Buthus occitanus tunetanus

C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 23-Aug-1996

C:Accession: A01746

R:Martin, M.F.; Rochat, H.

Toxicon 22, 279-291, 1984

A:Title: Purification of thirteen toxins active on mice from the venom of the North Afric

A:Reference number: A94316; MUID:84224814

A:Accession: A01746

A:Molecule type: protein

A:Residues: 1-65 <MAR>

C:Superfamily: scorpion neurotoxin

C:Keywords: neurotoxin; venom

F:12-63,16-36,22-46,26-48/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 20.5%; Score 69.5; DB 1; Length 65;

Matches 13; Conservative 10; Mismatches 22; Indels 5; Gaps 2;

QY 17 LEDGIVEAGFCGCPFNAGK---CHRHCKSIIRRRGGFCR--GTFRTTCVCYR 61

Db 1 LKDGIVDDRNCTYFCGTNAYCNBECYKLGESYCCQWVGYNACWCYK 50

RESULT 10

S11156

PSA10 protein - cowpea

C:Species: Vigna unguiculata (cowpea)

C:Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jun-2000

C:Accession: S11156

R:Ishibashi, N.; Yamauchi, D.; Minamikawa, T.

Plant Mol. Biol. 15, 59-64, 1990

A:Title: Stored mRNA in cotyledons of Vigna unguiculata seeds: nucleotide sequence of

A:Reference number: S11156; MUID:91355865

A:Accession: S11156

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-75 <ISH>

A:Cross-references: EMBL:X16877; NID:g22075; PID:g22076

C:Superfamily: gamma-thionin

Query Match 20.4%; Score 69; DB 2; Length 75;

Best Local Similarity 33.3%; Pred. No. 0.84;

Matches 24; Conservative 6; Mismatches 28; Indels 14; Gaps 4;

QY 2 KSIA-IIFIVLVAFPCILE-----DGIVEAGFCGCPFNAGKCHRHCKSIIR--RGGFC 49

Db 4 KSIAGLCFLFLVFAQEVVVOSEAKTCLNLDVTYRGCFCTTGSCDDHCKNKEHLLSGRC 63

QY 50 RGTFRTTCVCYR 61

Db 64 RDDVR--CWCTR 73

RESULT 11

T06599

disease resistance response protein 230 - garden pea

C:Species: Pisum sativum (garden pea)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jun-2000

C:Accession: T06599

R:Chiang, C.C.; Hadwiger, L.A.

Mol. Plant Microbe Interact. 4, 324-331, 1991

A:Title: The Fusarium solani-induced expression of a pea gene family encoding high cy

A:Reference number: Z15787; MUID:92190628

A:Accession: T06599

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-72 <CHI>

A:Cross-references: EMBL:L01578; NID:g169073; PIDN:AAA79117.1; PID:g169074

A:Experimental source: cv. Alaska

C:Genetics:

A:Gene: DRR230-a

C:Superfamily: gamma-thionin

Query Match

Best Local Similarity 20.2%; Score 68.5; DB 2; Length 72;

Matches 19; Conservative 11; Mismatches 24; Indels 13; Gaps 3;

QY 2 KSTATIFIVLVAFPCILEDGIVEAGFCGCPFNAGK-----CHRHCKSIIRRG---GFCR 50

Db 4 KSLACLFLLLVFAQEIWVSEANTCENLAGSKGVCFGGCDRHCR--QEGATSGRCR 61

QY 51 GTFRTTC 57

Db 62 DDFRCWC 68

RESULT 12

A31792

sapecin precursor - flesh fly (Sarcophaga peregrina)

2

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:43:24 ; Search time 25.63 Seconds
(without alignments)
411.732 Million cell updates/sec

Title: US-09-829-481-4
Perfect score: 339
Sequence: 1 MKSIAIFIVLVAFCILEDG.....IRRGGFCRGTRFTTCVCYR 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL19.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp-invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phase.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp-unclassified.*
- 15: sp-rviro.*
- 16: sp-bacteriap.*
- 17: sp-archaeap.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	154.5	45.6	81	Q9Y0A9	Q9Y0A9 mytilus gal
2	152	44.8	82	Q9U6U0	Q9U6U0 mytilus gal
3	146.5	43.2	73	Q9BLJ3	Q9BLJ3 ornithodor
4	140	41.3	73	Q9BLJ4	Q9BLJ4 ornithodor
5	108	31.9	57	Q9Y0B0	Q9Y0B0 mytilus gal
6	90	26.5	91	Q9N661	Q9N661 mesobuthus
7	82.5	24.3	85	Q9NJC7	Q9NJC7 mesobuthus
8	79.5	23.5	79	Q9G049	Q9G049 aegydes rhi
9	77	22.7	57	Q9G0Y6	Q9G0Y6 aedes aegypt
10	75.5	22.3	85	Q9NJP8	Q9NJP8 mesobuthus
11	75.5	22.3	85	Q9GYX2	Q9GYX2 mesobuthus
12	71.5	21.1	85	Q9BLM4	Q9BLM4 anodonta
13	70	20.6	102	Q61721	Q61721 anopheles g
14	70	20.6	103	Q9Y0B1	Q9Y0B1 mytilus gal
15	69.5	20.5	72	Q9FR81	Q9FR81 pisum sativ
16	68.5	20.2	85	Q9GNG8	Q9GNG8 mesobuthus

17	67.5	19.9	79	10	Q39807	Q39807 glycine max
18	66.5	19.6	85	5	Q9GUA7	Q9GUA7 mesobuthus
19	66.5	19.6	85	5	Q9SP69	Q9SP69 mesobuthus
20	65.5	19.3	195	11	Q91VZ7	Q91VZ7 mus musculu
21	64	18.9	96	5	Q77217	Q77217 aedes aegypt
22	64	18.9	98	5	P91793	P91793 aedes aegypt
23	64	18.9	98	5	Q963E9	Q963E9 aedes aegypt
24	63.5	18.7	81	10	Q9XGD9	Q9XGD9 zea mays (m
25	63.5	18.7	83	5	Q9BK52	Q9BK52 acalolepta
26	63	18.6	73	5	Q9FQ14	Q9FQ14 citrus para
27	62	18.3	87	10	Q948T4	Q948T4 pyrus pyrif
28	62	18.3	490	16	Q9PKG3	Q9PKG3 chlamydia m
29	62	18.3	990	13	Q91803	Q91803 xenopus lae
30	62	18.3	4599	11	Q9YI18	Q9YI18 mus musculu
31	61.5	18.1	85	5	Q9Y0B8	Q9Y0B8 mesobuthus
32	61.5	18.1	154	12	Q84631	Q84631 paramacium
33	61	18.0	39	5	P82380	P82380 stomoxys ca
34	61	18.0	643	11	Q9ET61	Q9ET61 rattus norv
35	61	18.0	643	11	Q9JIZ6	Q9JIZ6 rattus norv
36	61	18.0	644	11	O89103	O89103 mus musculu
37	60.5	17.8	81	5	Q94435	Q94435 centruride
38	60.5	17.8	85	5	Q9YX87	Q9YX87 mesobuthus
39	60.5	17.8	85	5	Q95WX6	Q95WX6 mesobuthus
40	60.5	17.8	126	5	O45448	O45448 caenorhabdi
41	60.5	17.8	652	4	Q9NFP3	Q9NFP3 homo sapien
42	60.5	17.8	652	4	O00274	O00274 homo sapien
43	60	17.7	85	5	Q9YIU3	Q9YIU3 mesobuthus
44	60	17.7	85	5	Q9UAC8	Q9UAC8 mesobuthus
45	60	17.7	142	4	Q9NZK7	Q9NZK7 homo sapien

ALIGNMENTS

RESULT 1

Q9Y0A9	PRELIMINARY;	PRT;	81 AA.
ID Q9Y0A9;			
AC 01-NOV-1999 (TReMBLrel. 12, Created)			
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)			
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)			
DE DEFENSIN MGD-2 PRECURSOR.			
OS Mytilus galloprovincialis (Mediterranean mussel).			
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;			
OC Mytiloidae; Mytilidae; Mytilus.			
OX NCBI_TaxID=29158;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=HEMOCYTE;			
RX MEDLINE=20033584; PubMed=10564642;			
RA Mitta G., Vandenbulcke F., Hubert F., Roch P.;			
RT "Mussel defensins are synthesized and processed in granulocytes then released into the plasma after bacterial challenge.";			
RL J. Cell Sci. 112:4233-4242(1999).			
CC -!- SUBCELLULAR LOCATION: SECRETED.			
CC -!- MISCELLANEOUS: THE PI OF THIS PROTEIN IS 8.92.			
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.			
DR EMBL; AF162338; AAD45118.1;			
DR InterPro; IPR001542; Arthro_defensin.			
DR InterPro; IPR003614; Knott.			
DR Pfam; PF01097; Arthro_defensin; 1.			
DR SMART; SM00505; Knott1; 1.			
DR PROSITE; PS00425; ARTHROPOD_DEFENSIN; UNKNOWN_1.			
KW Defensin; Antibiotic; Fungicide; Amidation; Signal.			
FT SIGNAL 1 21 POTENTIAL.			
FT PROPEP 22 60 POTENTIAL.			
FT CHAIN 61 81 DEFENSIN MGD-2.			
FT MOD_RES 60 60 AMIDATION (POTENTIAL).			
FT DISULFID 25 46 BY SIMILARITY.			
FT DISULFID 31 54 BY SIMILARITY.			
FT DISULFID 35 56 BY SIMILARITY.			
FT DISULFID 42 59 BY SIMILARITY.			
FT DISULFID 42 59 BY SIMILARITY.			
SQ SEQUENCE 81 AA; 9087 MW; 0EEF7AB5CF6DCE53 CRC64;			

```

Query Match 45.6%; Score 154.5; DB 5; Length 81;
Best Local Similarity 51.7%; Pred. No. 1.3e-12;
Matches 30; Conservative 7; Mismatches 18; Indels 3; Gaps 3;

Qy 5 AIIIFIVLVAFCILEDGIVEAGFCPPFNAGCKHRHCKSIRRR-GGFCRGFTRTTCVCYR 61
| : : : | | : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 AAFVLLVVLGLIMTD-VATAGFCPPNYYA-CHQCKSIRGCGYGAGWFLRCLTCYR 58

RESULT 2
Q9BLJ3 ID Q9BLJ3 PRELIMINARY; PRT; 82 AA.
AC Q9BLJ3
DT 01-JUN-2001 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ANTIMICROBIAL PEPTIDE MGD2B.
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakajima Y.;
RT "CDNA CLONING OF TICK DEFENSIN.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RT gene structure and expression analysis.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177539; AAD52660.1; -.
DR InterPro; IPR001542; Arthro_defensin.
DR InterPro; IPR001545; Glyco_hormone_beta.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.
SQ SEQUENCE 82 AA; 9307 MW; 6D8F0039FE675ECE CRC64;

Query Match 44.8%; Score 152; DB 5; Length 82;
Best Local Similarity 48.3%; Pred. No. 2.9e-12;
Matches 28; Conservative 8; Mismatches 20; Indels 2; Gaps 2;

Qy 5 AIIIFIVLVAFCILEDGIVEAGFCPPFNAGCKHRHCKSIRRR-GGFCRGFTRTTCVCYR 61
| : : : | | : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 AAFVLLVVLGLIMMDVATAGFCPPNYYA-CHQCKSIRGCGYGAGWFLRCLTCYR 59

RESULT 3
Q9BLJ3 ID Q9BLJ3 PRELIMINARY; PRT; 73 AA.
AC Q9BLJ3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE DEFENSIN A.
GN OMDEF-A.
OC Ornithodoros moubata (Soft tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Argasidae; Ornithodoros.
OX NCBI_TaxID=6938;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakajima Y.;
RT "CDNA CLONING OF TICK DEFENSIN.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB041816; BAB41028.1; -.
DR InterPro; IPR001542; Arthro_defensin.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.
SQ SEQUENCE 73 AA; 8191 MW; D94B4AC5BDA89FFD CRC64;

Query Match 43.2%; Score 146.5; DB 5; Length 73;
Best Local Similarity 40.3%; Pred. No. 9.5e-07;
Matches 27; Conservative 11; Mismatches 19; Indels 10; Gaps 3;

Qy 4 IAIIFIVLVAFCILE--DGI-----VEAGFCPPFNAGCKHRHCKSIR-RRGGFCRGTF 53
| : : : | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7 VAIIVLAVATMAQAEVDDVEQSVPRVRRGYGCPFNQYQCHSHCKRGIRYKGYGTCGRF 66

Qy 54 RTTCVCY 60
| : : : | | | | |
Db 67 KQCKCY 73

RESULT 5
Q9YOB0 ID Q9YOB0 PRELIMINARY; PRT; 57 AA.
AC Q9YOB0
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MGD1 ANTIMICROBIAL PEPTIDE (FRAGMENT).
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE FROM N.A.
RA Mitta G.; Hubert F.; Roch P.;
RT "Defensins: involvement in mussel defense.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162337; AAD45117.1; -.
DR InterPro; IPR001542; Arthro_defensin.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
FT NON_TER 1
SQ SEQUENCE 57 AA; 6658 MW; 7B86B8623C979326 CRC64;

```

DR	HSP; P01484; 1AHO.
DR	InterPro; IPR003614; KntI.
DR	InterPro; IPR002061; Scorpion_toxin.
DR	InterPro; IPR001219; Toxin.
DR	PRINTS; PR00284; TOXIN.
DR	PDom; PD000908; Scorpion_toxin; 1.
DR	SMART; SM00505; KntI; 1.
KW	Signal.
FT	SIGNAL 1 19 POTENTIAL.
FT	CHAIN 20 85 ALPHA TOXIN 2.
SO	SEQUENCE 85 AA; 9524 MW; 565390HB6E71806E CRC64;
Query Match 24.3%; Score 82.5; DB 5; Length 85;	
Best Local Similarity 25.8%; Pred. No. 0.0027;	
Matches	17; Conservative 13; Mismatches 31; Indels 5; Gaps
QY	1 MKSIATIFVLVAFCEIDGIVEAGFGCPFNAGK---CHRRCKSIRRRGGFCR--CTFRT 55
DB	I : :: : :: : :: : :: : :: : :: : :: : :: : :: :
DB	4 MVIIALLVMVTGVESVDGYIADDRNCPYFCGRNAYCDGECKKNRAESYGQWASKYGN 63
QY	56 TCVCYR 61
DB	I : :: : :: : :: : :: : :: : :: : :: : :: :
DB	64 ACWCYK 69
RESULT 8	
R96049	PRELIMINARY; PRT; 79 AA.
ID	O96049 PRELIMINARY; PRT; 79 AA.
AC	O96049; 01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	DEFENSIN PRECURSOR.
OS	Oryctes rhinoceros (Coconut rhinoceros beetle).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC	Scarabaeiformia; Scarabaeidae; Dynastinae; Oryctes.
OX	NCBI_TaxID=72550;
RN	[1]
RP	SEQUENCE FROM N.A., SEQUENCE OF 35-58, AND CHARACTERIZATION.
RC	TISSUE=FAT BODY, AND HEMOLYPH;
RC	PubMed=10561605;
RA	Ishibashi J., Yang J., Saido-Sakanaka H., Nakazawa H., Yamamoto M.,
RA	Yamakawa M.;
FT	"Purification, cDNA cloning and modification of a defensin from the
FT	coconut rhinoceros beetle, Oryctes rhinoceros.";
FL	Eur. J. Biochem. 266:616-623(1999).
CC	-I- FUNCTION: SHOWS VERY STRONG ACTIVITY AGAINST S.AUREUS.
CC	-I- TISSUE SPECIFICITY: FAT BODY, HEMOCYTES, MALPIGHIAN TUBULES AND MIDGUT.
CC	-I- INDUCTION: BY BACTERIAL INFECTION IN FAT BODY, HEMOCYTES AND MALPIGHIAN TUBULES.
CC	-I- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR	EMBL; AB011245; BAA36401.1; -.
DR	HSSP; P10891; LICA.
DR	InterPro; IPR001542; Arthro_defensin.
DR	InterPro; IPR003614; KntI.
DR	Pfam; PF01097; Arthro_defensin; 1.
DR	PRINTS; PR00271; DEFENSIN.
DR	SMART; SM00505; KntI; 1.
DR	PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW	Insect immunity; Antibiotic; Defensin; Signal.
FT	SIGNAL 1 19 POTENTIAL.
FT	PROPEP 20 34
FT	CHAIN 35 79 DEFENSIN.
FT	DISULFD 39 70 BY SIMILARITY.
FT	DISULFD 56 75 BY SIMILARITY.
FT	DISULFD 60 77 BY SIMILARITY.
SO	SEQUENCE 79 AA; 8394 MW; 8D8A3EF8CDD3A83 CRC64;
Query Match 23.5%; Score 79.5; DB 5; Length 79;	

```

Best Local Similarity 28.8%; Pred. No. 0.0062;
Matches 23; Conservative 7; Mismatches 21; Indels 29; Gaps 4;

Qy 6 IIFIVLVAFCI-----LEDGIVE-----AGFGCPNAGKCHRHCKS 41
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 5 IVFAFIVAMCTAHLAAPAFALEASVIRKRLTCLLSFEAKGFAA--NHSICAHAHCLA 62

Qy 42 IRRRGGFCRGTFRTTCVCYR 61
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 63 IGRKGAGCQ---NGVCVCRR 79

RESULT 9
ID Q9GYU6 PRELIMINARY; PRT; 57 AA.
AC Q9GYU6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DEFENSIN (FRAGMENT).
OS Aedes albopictus (Forest day mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Aedes.
OX NCBI_TaxID=7160;
RN [1]
RP SEQUENCE FROM N.A.
RA Melinda P.M., Kostas B., Scott O.L.;
RT "Melbachia neither induces nor suppresses transcripts encoding
   antinicrobial peptides."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY005473; AAF97983.1; -
DR HSP; P10891; IICA.
DR InterPro; IPR001542; Arthro_defensin.
DR InterPro; IPR003614; Knot1.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PRINTS; PR00271; DEFENSIN.
DR SMART; SM00505; Knot1; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
FT NON_TER 1
FT SEQUENCE 57 AA; 6214 MW; A9F9DF5A8CA86A7 CRC64;

Query Match 22.7%; Score 77; DB 5; Length 57;
Best Local Similarity 44.4%; Pred. No. 0.0095;
Matches 16; Conservative 5; Mismatches 11; Indels 4; Gaps 2;

Qy 24 AGFGCPFNAGKCHRHCKSIRRRGGFCRGTFRTTCVC 59
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 26 SGFGVGDSA--CAAHCIARRNRGGYCNA--KTVCVC 57

RESULT 10
Q9NJUP8 PRELIMINARY; PRT; 85 AA.
AC Q9NJUP8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEUROTOXIN TX11P.
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Buthus.
OX NCBI_TaxID=34649;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-VENOM GLAND;
RC Zhu S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF132974; AAF31295.1; -
DR HSP; P01484; IAHO.
DR InterPro; IPR003614; Knot1.

```

```

DR InterPro; IPR002061; Scorpion_toxin.
DR InterPro; IPR001219; Toxin.
DR Pfam; PF00537; toxin_3; 1.
DR PRINTS; PR00284; TOXIN.
DR ProDom; PD000908; Scorpion_toxin; 1.
DR SMART; SM00505; Knot1; 1.
SQ SEQUENCE 85 AA; 9588 MW; 416CB3D72A8927ED CRC64;

Query Match 22.3%; Score 75.5; DB 5; Length 85;
Best Local Similarity 26.6%; Pred. No. 0.022;
Matches 17; Conservative 13; Mismatches 27; Indels 7; Gaps 3;

Qy 3 SIALLIFVLVAFCILEDGIVEAGFGCPFNAGK---CHRHCKSIRRRGGFCR--GTFRTTC 57
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 8 SLALLMTGVE--SVKDGVIADDRNCPYFCGRNAYCDGCECKKNRAESGYCQWASKYGNAC 65

Qy 58 VCVR 61
   ||:
Db 66 WCYK 69

RESULT 11
Q9GYX2 PRELIMINARY; PRT; 85 AA.
ID Q9GYX2;
AC Q9GYX2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA_TOXIN 1 PRECURSOR.
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Buthus.
OX NCBI_TaxID=34649;
RN [1]
RP SEQUENCE FROM N.A.
RA Ye J., Chen J., Zuo X., Ji Y.;
RT "Cloning and characterization of cDNA sequences encoding two novel
   alpha toxin precursors from the Chinese scorpion Buthus martensii
   Karsch."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF288607; AAG00580.1; -
DR HSP; P17728; ILQI.
DR InterPro; IPR003614; Knot1.
DR InterPro; IPR002061; Scorpion_toxin.
DR InterPro; IPR001219; Toxin.
DR PRINTS; PR00284; TOXIN.
DR ProDom; PD000908; Scorpion_toxin; 1.
DR SMART; SM00505; Knot1; 1.
KW Signal.
FT CHAIN 1 19 POTENTIAL.
FT CHAIN 20 85 ALPHA_TOXIN 1.
SQ SEQUENCE 85 AA; 9366 MW; 94ED463D6193FFEF CRC64;

Query Match 22.3%; Score 75.5; DB 5; Length 85;
Best Local Similarity 26.6%; Pred. No. 0.022;
Matches 17; Conservative 12; Mismatches 28; Indels 7; Gaps 3;

Qy 3 SIALLIFVLVAFCILEDGIVEAGFGCPFNAGK---CHRHCKSIRRRGGFCR--GTFRTTC 57
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 8 SLALLMTGVE--SVRDGYIADDRNCPYFCGRNAYCDGCECKKNRAESGYCQWASKYGNAC 65

Qy 58 VCVR 61
   ||:
Db 66 WCYK 69

RESULT 12
Q9BLM4 PRELIMINARY; PRT; 85 AA.
ID Q9BLM4;
AC Q9BLM4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```


DT	01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT	01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE	PUTATIVE TOXIN PRECURSOR.
OS	Androctonus australis (Sahara scorpion).
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC	Buthoidea; Buthidae; Androctonus.
OX	NCBI_taxid=6858;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21211344; PubMed=11311249;
RA	Ceard B., Martin-Eauclaire M.F., Bougis P.E.;
RT	"Evidence for a position-specific deletion as an evolutionary link
RT	between long- and short-chain scorpion toxins.";
RL	FEMS Lett. 494:246-248(2001).
DR	EMBL; AJ308440; CAC37321.1; -.
DR	HSSP; P01484; IAHO.
DR	InterPro; IPR003614; Knot1.
DR	InterPro; IPR002061; Scorpion_toxin.
DR	InterPro; IPR001219; Toxin.
DR	PRINTS; PR00284; TOXIN.
DR	ProDom; PD000908; Scorpion_toxin; 1.
DR	SMART; SM00505; Knot1; 1.
KW	Signal.
FT	SIGNAL
FT	CHAIN
FT	SEQUENCE 85 AA; 9523 MW; 4059A69D80E4F090 CRC64;
QY	Query Match
DB	Best Local Similarity 21.1%; Score 71.5; DB 5; Length 85;
DB	Matches 18; Conservative 12; Mismatches 29; Indels 7; Gap 5
QY	1 MKSIAIIFIVLVAFCEILEDGIVEAGFCGPNACK---CHRCKSIARRRGFCR--GTFRF 5
DB	6 MISLALLEMTGVE--SKKDGVIYDKNKCTFFCGRNAYCNDEKKKGAEISGYCQWASPYGN 6
QY	56 TCVCYR 61
DB	ACYCYK 59
RESULT 13	
ID	PRELIMINARY; PRT; 102 AA.
AC	061721
DC	061721;
AT	01-AUG-1998 (TReMBLrel. 07, Created)
DT	01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT	01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE	DEFENSIN.
GN	DEF.
OS	Anopheles gambiae (African malaria mosquito).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC	Anopheles.
OX	NCBI_taxid=7165;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=KWA;
RA	Eggleston P., Lu W., Zhao Y.;
RL	Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RL	EMBL; AF063402; AAC18575.1; -.
DR	HSSP; p10891; LIICA.
DR	InterPro; IPR001542; Arthro_defensin.
DR	InterPro; IPR003614; Knot1.
DR	Pfam; PF01097; Arthro_defensin; 1.
DR	PRINTS; PR00271; DEFENSIN.
DR	SMART; SM00505; Knot1; 1.
DR	PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
QY	SEQUENCE 102 AA; 10681 MW; 6288289B1620CD0C CRC64;
Query Match	20.6%; Score 70; DB 5; Length 102;
Best Local Similarity	41.7%; Pred. No. 0.13;

